

NAME

pds-to-cdaw -- converting PDS label/table data sets to CDAWeb CDF files

DESCRIPTION

This document describes the use of the *pds-to-cdf* program to translate PDS label/table files into CDF files which meet the special requirements of the CDAWeb web site.

GENERATING CDAWEB FILES

CDAWeb is a web site at <http://cdaweb.gsfc.nasa.gov/> which hosts a collection of CDF-formatted data sets, and provides tools to search the collection and view the data in various representations. The emphasis of CDAWeb is on plotting a selection of observation values as a function of their observation times.

The *pds-to-cdf* conversion program provides several options which can be used to generate CDF files that meet the special requirements imposed by the CDAWeb software: **--cdaweb**, **--template**, and **--epoch**. The **--template** option has some general-purpose utility, while the **--cdaweb** and **--epoch** options are very CDAWeb-specific. The generation of a CDAWeb-compliant CDF file will require the use of all three of these options.

CDAWEB MISCELLANEA

The **--cdaweb** option activates several CDAWeb-specific actions to be performed in the generation of the CDF file. The use of the **--cdaweb** option also requires the use of both the **--template** and the **--epoch** options, described subsequently.

SELECTING THE 'EPOCH' VALUES

The CDAWeb web site displays tabular data as a series of observation values plotted against a common set of observation times. These observation times are referred to as the "epochs". With this nomenclature, a data set could easily contain thousands of observation times -- "epochs" -- with each epochal period spanning all of a second or two.

The **--epoch** option is used both to select the column from the PDS table file that is to be used to form these epoch values, and to select the conversion algorithm to be used. The available options are enumerated in the **pds-to-cdf(1)** man page. The time values are converted into values of the CDF_EPOCH data type, which uses a floating point internal representation of these observation times. The resulting CDF_EPOCH values are stored in the "Epoch" variable of the generated CDF file.

SUPPLYING GLOBAL ATTRIBUTES

The CDAWeb web site requires that all CDF files include several CDF global attributes that describe each data set. These CDF global attributes are supplied to the *pds-to-cdf* program by creating a CDF file containing these attribute values and passing the name of this "template" CDF file to the **--template** option of the *pds-to-cdf* program. Such a CDF file can be produced by writing a CDF skeleton file and processing it with the **skeletoncdf(1)** program to generate a suitable template CDF file. In addition to the CDAWeb-mandated attributes, this same mechanism can be used to add any collection of CDF global attributes to a generated CDF output file.

The CDAWeb requirements for these global attributes are fully described in http://spdf.gsfc.nasa.gov/sp_use_of_cdf.html. In brief, however:

Logical_source_description

Supplies a single-line brief description of the data set.

Project, PI_name, PI_affiliation

Free-form single-value fields which name the project, its principal investigator, and the principal investigator's institutional affiliation.

Mission_group, Instrument_type

Used by the CDAWeb software to categorize the data set. Each attribute must take on one of several values established by the CDAWeb project.

The next few attributes require two-part values, in which two related fields are stored in a single global attribute, separated by the "greater-than" symbol ">".

Discipline

This attribute must supply both a general category and a more specific subcategory for the science discipline of the data set, as for example "Space Physics>Interplanetary Studies".

Source_name, Data_type, Descriptor

Each of these attributes must supply both a single-word and a brief description of the attribute in question, as for example, a **Descriptor** value of "1hr>One hour".

The first element in each of these last three attribute values are used to form the **Logical_source** attribute value by joining them together, this time using an underscore character "_" to separate the three fields. Another attribute, **File_naming_convention** is set to a value of "source_datatype_descriptor" to record this ordering.

Finally, a **Logical_file_id** attribute value is generated by combining the **Logical_source** value, a date field, and a version number. The date field is the four-digit year, two-digit month, and two digit day number of the start of the data set, unless the data set begins after noon, in which case the following day is used to form the date field. A version number of "1" is used unless overridden by supplying a different value for the **Data_version** attribute value. This **Logical_file_id** attribute value is also used as the file name of the generated CDF file.

OTHER ATTRIBUTES

There are several other global attributes and variable attributes included to satisfy the requirements of the CDAWeb web site. Since these attributes can in most cases be derived directly from information in the PDS source files, they are included even when the **--cdaweb** option is not in effect.

An overall description of the dataset is taken from the PDS "DESCRIPTION" field of either the table element or the root element of the PDS label file, and stored in the somewhat less than descriptively named **TEXT** global CDF attribute.

The remaining attributes are associated with individual CDF variables, which contain the contents of the PDS table columns. Four of these variable attributes (**VAR_NOTES**, **CATDESC**, **FIELDNAM** and **LABLAXIS**) are based on the "DESCRIPTION" field of each label file "COLUMN" element.

VAR_NOTES

The entire "DESCRIPTION" field taken directly from the PDS "COLUMN" element

CATDESC

The "DESCRIPTION" field with newline and carriage return characters removed, and with contiguous runs of white-space characters replaced by a single space character.

FIELDNAM, LABLAXIS

The initial portion of the "DESCRIPTION" field, truncated at a word boundary prior to about 40 characters.

The next three variable attributes (**FILLVAL**, **VALIDMIN**, and **VALIDMAX**) describe the range of values encountered in a column.

FILLVAL

This variable attribute is set to -2147483648 (-2^{31}) for integer valued data and to -1.0e31 for real valued data. The value of the **FILLVAL** attribute is used as a replacement value for any data point value that matches the PDS "MISSING_VALUE" value for this column.

VALIDMIN , VALIDMAX

The minimum and maximum column value, exclusive of any datum having the value specified in the "MISSING_CONSTANT" entry of the PDS column element.

The remaining variable attributes have fairly unremarkable derivations.

FORMAT and UNIT

The **FORMAT** and **UNIT** attributes are taken directly from the PDS column elements of the same names.

DEPEND_0

The **DEPEND_0** variable attribute is always set to a value of "Epoch".

VAR_TYPE

The **VAR_TYPE** variable attribute is set to a value of "support_data" for the **Epoch** CDF variable, and to a value of "data" for all other attributes.

DISPLAY_TYPE

The **DISPLAY_TYPE** variable attribute is set to a value of "time_series" for all non-**Epoch** variables, and left unset for the **Epoch** variable.

SEE ALSO

pds-to-cdf(1), **skeletoncdf(1)**